Speech Intelligibility with Acoustic and Contact Microphones

Barbara Acker-Mills, Ph.D., Adrianus Houtsma, Ph.D., and William Ahroon, Ph.D.
U.S. Army Aeromedical Research Laboratory
P.O. Box 620577
Fort Rucker, AL 36362-0577

barbara.acker@us.army.mil

SUMMARY

Speech intelligibility of signals obtained with an acoustic microphone and three types of vibration-driven contact microphones was assessed using the Diagnostic Rhyme Test (DRT). Stimulus words were recorded digitally in a reverberant chamber with no noise and with ambient broadband noise intensity at 106 dB(A). Listeners completed the DRT task in the same settings, thus simulating typical environments of a rotary-wing aircraft. Results show that speech intelligibility is significantly worse for the contact microphones than for the acoustic microphone, particularly in noisy environments, and some consonant types are affected more than others. Therefore, contact microphones are not recommended for use in any situation where fast and accurate speech intelligibility is essential.

INTRODUCTION

The ability to communicate effectively is of paramount importance in the rotary-wing aircraft environment. Effective communication leads to increased aircrew safety and performance and contributes to successful mission completion. As communication is degraded, mission capability is reduced and the safety of the aircrew is compromised. A communication system involves at least one “talker” (sender), one “listener” (receiver), and any equipment used to augment or transmit information. Most research focuses on the listener and on devices that can increase the speech-to-noise ratio and reduce noise-induced hearing loss.

While the standard U.S. Army aviator helmet, the HGU-56/P Aircrew Integrated Helmet System, is designed primarily for impact protection, it also includes a set of headphones mounted in sound-attenuating earcups, a noise-canceling acoustic microphone (the boom microphone), and an optional Communications Earplug (CEP). The boom microphone, positioned close to the speaker’s mouth, consists of two transducer elements that are faced in opposite directions and are wired out of phase. Thus, a diffuse noise field yields a small residual output signal while directed speech sound yields a much larger differential output. Thus, by limiting the impact of ambient noise, the helmet improves the speech-to-noise ratio (for the listener) and also protects the crew from noise-induced hearing loss (for the listener and talker). A large corpus of information exists on the listener component of the communication system, but very little research assesses problems at the talker level.

Although the noise-cancelling boom microphone in the HGU-56/P works well if positioned and used properly, improper microphone use and noise conditions may impair performance. Additionally, an open microphone, in contrast to the usual “keyed” microphone, often is necessary in situations that require use of both hands (e.g., a crew chief operating a hoist may need to use both hands on the hoist control and cable). Indeed, there

Acker-Mills, B.; Houtsma, A.; Ahroon, W. (2005) Speech Intelligibility with Acoustic and Contact Microphones. In *New Directions for Improving Audio Effectiveness* (pp. 7-1 – 7-14). Meeting Proceedings RTO-MP-HFM-123, Paper 7. Neuilly-sur-Seine, France: RTO. Available from: http://www.rto.nato.int/abstracts.aps.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **01 APR 2005**	2. REPORT TYPE **N/A**	3. DATES COVERED **-**

4. TITLE AND SUBTITLE **Speech Intelligibility with Acoustic and Contact Microphones**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **U.S. Army Aeromedical Research Laboratory P.O. Box 620577 Fort Rucker, AL 36362-0577**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
See also ADM001856, New Directions for Improving Audio Effectiveness (Nouvelles orientations pour l'amelioration des techniques audio)., The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**UU**	**14**	

are some aircraft environments in which an open microphone is the normal operating condition (e.g., in the British Army). Furthermore, noise conditions may be encountered (e.g., air movement from an open window or door) for which the normal noise-canceling boom microphone was not designed to minimize. These problems surrounding boom microphones are not new. During World War II, acoustic microphones faced similar noise issues, signals were very noisy, and microphones were impractical for manual operations or for tasks requiring excessive head motion (6). Throat microphones were developed in response to these problems and used during the later years of WWII.

Throat microphones are designed to pick up (transduce) the vibrations of the vocal apparatus at the throat instead of the vibrations of air molecules at the mouth. Since microphones convert sound signals (spoken words) into electrical signals to be transmitted into the communication system, boom microphones must necessarily also transmit any ambient noise around the talker's mouth. Removing this ambient noise from the communication signal should improve the signal-to-noise ratio for the listener. Thus, by virtue of being relatively insensitive to airborne sound, throat microphones, as compared to boom microphones, should produce a signal that contains less noise.

There are a variety of bone-conduction communication systems, but all operate on the same principle. The microphone is placed somewhere on the skull and is sensitive to internal vibrations created by the production of speech waves that travel through the facial and skull bones to the microphone. The current study used two different communication systems that included both a microphone and a loudspeaker. The head-gear system consisted of a microphone that was in contact with the top of the skull, and bone-conduction speakers that were in contact with both sides of the forehead and with both upper jaw bones. The other system contained a bone-conduction microphone in an earpiece, but the speaker was a regular acoustic speaker. The whole unit fit in the ear canal and was worn under a helmet.

During the development of throat and other contact microphones, there has been little systematic evaluation of speech intelligibility in noise using these devices, and the few existing studies contain conflicting results. Snidecor, Rehman, and Washburn (10) explored vowel intelligibility with contact microphones located on different areas of the head and neck. Stimuli were recorded in quiet and presented over headphones to listeners who also were in a quiet environment. One group of listeners rated intelligibility and another group of listeners gave quality judgments of the vowels. Contact microphone locations at the forehead, mastoid, and larynx were highest in intelligibility and quality ratings. While somewhat informative, this study does not use objective speech intelligibility measures, and with all recording and listening completed in a quiet environment, does not assess how contact microphones might function in noise.

Oyer (9) evaluated intelligibility of words recorded simultaneously with an acoustic microphone placed at the mouth and another acoustic microphone placed in the ear canal. The microphone placed in the ear canal was intended to pick up vibrations created by speech signals and transmitted through the skull to the ear canal. Air traffic control words in carrier phrases were recorded in quiet and mixed with 74 dB white noise prior to being presented to subjects over standard headphones. The signal-to-noise ratio was manipulated by attenuating the speech signal (-12, -15, and -18 dB). Results revealed a microphone × signal-to-noise ratio interaction, where speech intelligibility decreased for both microphones as the signal-to-noise ratio decreased, but the decrement was less for the ear microphone. Although not part of the formal study, it was reported anecdotally that simultaneous presentation of the acoustic and ear microphone stimuli resulted in very good speech intelligibility.

A study by Moser and Dreher (7) is most relevant to the current research project. They used a noise-canceling acoustic microphone, an ear microphone, and a bone conduction microphone placed on the forehead to record

the Phonetically Balanced (PB) word lists (developed by Egan (3) and later standardized by the American National Standards Institute (ANSI) S3.2-1989 (1)). The words were recorded by pilots in two different transport aircraft. Ambient noise in the KC-97 aircraft was measured at 97 dB(C) and was measured at 106 dB(C) in the C-124 aircraft. Listeners were in a quiet environment and stimuli were presented at 77 dB(C) over PDR-8 acoustic headphones. Listeners became familiar with the two different microphone stimuli by listening to a paragraph followed by operational instructions. They then completed the PB task which consisted of writing the word that was presented through headphones. All microphone transmissions were judged as “acceptable” during the familiarization phase, but PB results were better with the acoustic microphone than either the bone or ear microphone in both aircraft environments. Moser and Dreher (7) also conducted an informal evaluation of bone and ear receivers (not microphones) in an aircraft and found ear receivers to be rated as “excellent.” The bone conduction receivers placed on the mastoid, however, were considered “not acceptable” if the ears were not shielded from the aircraft noise.

In view of the sparse experimental research concerning the original contact microphones, it is surprising that very few studies address speech intelligibility using modern contact microphones. In fact, a thorough literature search found only one peer-reviewed paper that mentioned speech intelligibility secondary to a study of temporary threshold shifts (4). A few recent conference presentations have discussed bone-conduction communication systems (5), but these studies have not yet been published, nor did the listening conditions approximate the noisy environment of rotary-wing aircraft.

Even in the absence of speech intelligibility data, contact microphones are marketed to law enforcement agencies (e.g., Los Angeles Police Department SWAT), fire departments, and to a variety of users for applications that require special environmental controls (respirators, hazardous material suits, etc.) or extremely rugged construction (waterproof, dustproof). Several segments in the DoD are strong advocates of these devices, but systematic research evaluating the intelligibility of speech transmitted with the devices should be completed before recommendations can be made for use in military environments.

A direct comparison of the boom and contact microphones is important since each has its own unique strengths and weaknesses as related to transmitting a speech signal. Although contact microphones may increase the overall signal-to-noise ratio, they also may produce a speech signal that is less intelligible than that produced by boom microphones because of the lack of encoding some important speech components. Specifically, consonant sounds are produced with the articulators (the tongue, lips, jaw position, etc.). Because throat microphones pick up information before the level of the articulators, they should not be very effective in transmitting consonant sounds. However, throat microphones should effectively transmit vowel sounds (which are produced by the vocal cords). Boom microphones, on the other hand, are located close to the articulators and thus should efficiently transmit consonant sounds, but with the trade-off of also transmitting ambient noise, resulting in a lower signal-to-noise ratio.

Throat microphones make contact with the soft tissue of the throat and record information before the effects of the articulators (hard/soft palate, tongue, lips, and teeth) have been added. It is possible that bone-conduction microphones will be more effective, because the bones vibrate in response to vibrations of the vocal cords and the articulators. Thus, less ambient noise will be recorded than with an acoustic microphone, but more consonant and vowel information will be available than with the throat microphone.

The current study provides an objective, experimental evaluation of speech intelligibility for stimuli recorded using the HGU-56/P acoustic microphone, and commercially-available throat and bone-conduction microphones. The experimental conditions include realistic noise conditions and thus address the feasibility of use of these microphone options in rotary-wing aircraft.

EXPERIMENT 1: THROAT MICROPHONE

Method

Speech intelligibility was measured with the Diagnostic Rhyme Test (DRT) using procedures specified by ANSI S3.2-1989 (1). The test stimuli consists of six categories of consonants, with each category containing 16 word pairs that differ only in the initial consonant. The six consonant categories are voicing, nasality, sustention, sibilation, graveness, and compactness. These categories are based on acoustical properties of the consonants, not on place and manner of articulation. Table 1 describes the categories (8).

Table 1: Diagnostic Rhyme Test consonant category descriptions.

Category	Description	Example
Voicing (V)	Voiced (vocal cords vibrate) and voiceless consonants (no vibration).	bean (voiced) vs. peen (voiceless)
Nasality (N)	The initial nasal stops (voiced) are paired with their bilabial stop counterparts.	meat (nasal) vs. beat (voiced, bilabial stop)
Sustention (Sust)	No movement compared with movement of the articulators during production. Sustained speech instead of interrupted as with a stop consonant.	vee (sustained) vs. bee (interrupted)
Sibilation (Sib)	Fricatives accompanied by a hissing sound; produces aperiodicity in the high frequencies. Paired with non-hiss fricatives.	zee (hiss) vs. thee (no hiss)
Graveness (G)	Labial consonants (produced at lips) with energy focused in the lower frequencies. Paired with consonants produced further back in the mouth (alveolars, palatals, etc.).	weed (labial) vs. reed (alveolar)
Compactness (C)	Consonants produced with a concentration of energy in a narrow, central area of the spectrum. Paired with more spectrally diffuse consonants.	key (narrow) vs. tea (diffuse)

Stimuli

The 96 DRT stimuli were recorded in a reverberant chamber without background noise and with a background of spectrally-shaped broad-band noise intensity of 106 db(A). The latter simulates a UH-60A Black Hawk helicopter in straight-and-level flight at 120 knots indicated airspeed. A single-transducer LASH II throat microphone (based on the Thales Acoustics RA440 throat microphone) was used in conjunction with the HGU-56/P noise-canceling acoustic boom microphone to record the stimuli. The LASH II throat microphone is a small lightweight device with medium sensitivity (-47 dB re 1V/Pa) specifically designed for use in very high noise environments such as rotary-wing aircraft. The microphone has a frequency response of about 150 to 5000 Hz which is an improvement over throat microphones used in the past.

The male talker wore a throat microphone and an HGU-56/P helmet with the standard noise-canceling boom microphone (frequency sensitivity from about 200 Hz to 6000 Hz). The talker fastened the throat microphone at a comfortable position and pressure, which was measured at about 200 grams of force. (Thales Acoustics

does not provide specific directions for use of the throat microphone, except that it should fit comfortably without undue pressure.) The sound-attenuating earcups of the HGU-56/P and use of the CEP protected the talker from noise during the recording session.

Two separate analog-to-digital channels (16-bit, 40kHz sampling rate) were used to record stimuli simultaneously from the two separate microphones. Two separate sound files, one from the boom microphone and one from the throat microphone, were created for each stimulus. Stimuli were post-processed to ensure equivalent overall root mean square levels within each microphone type.

Several stimuli were selected randomly for spectral analysis. Because each of the target words contained different types of consonants that differ in intensity during natural speech, none of the signal-to-noise ratios were exactly the same. However, the signal-to-noise ratios for stimuli recorded from the throat microphone were always higher (approximately 10 dB) than for stimuli from the boom microphone.

Participants

Participants were Soldiers at Fort Rucker, Alabama, awaiting the Army Warrant Officer Course and flight training school. Eight males (mean age = 25) volunteered for the study. All but one volunteer had normal hearing as confirmed by recent physical exams or by audiograms performed at the U.S. Army Aeromedical Research Laboratory (USAARL) acoustics laboratory. One male (age 43) reported having tinnitus in his right ear. The study protocol was approved in advance by the USAARL Human Use Committee, and each participant provided written informed consent before participating.

Procedure

All testing took place in the USAARL-Acoustics Laboratory reverberant chamber, and stimulus presentation/response collection was coordinated using the Avaaz Experiment Generator and Controller software (2). The purpose of the study and experimental conditions were explained to the participant and then the participant was fitted with an HGU-56/P. The headphones in this helmet have a frequency range of about 200 Hz to 5000 Hz.

The DRT word-pairs were visually displayed on a computer monitor, followed by the target word presented in the earphones. Participants had 3 seconds in which to use a mouse to select the target word that was spoken. There was a 500 ms delay between the word-pairs appearing on the screen and the auditory stimulus presentation. The correct choice was displayed 500 ms after a response and lasted for 750 ms. If a participant did not respond in the allocated 3000 ms, an incorrect response was recorded. Trials were separated by 750 ms.

A 2 × 2 × 6 repeated-measures factorial design was used, with microphone type (boom, throat), noise (none, 106 dB(A)), and consonant category (see Table 1) as the independent variables. Each block contained 96 word-pair visual displays, with microphone and noise type held constant. Word-pairs, the order of words within a pair, and the word presented over headphones were randomized within and across blocks to avoid learning effects that would occur if there was any consistency across conditions. In addition, this procedure resulted in the consonant categories being completely randomized within blocks of trials so participants would not focus on any specific types of consonants (e.g., voicing, nasality, etc). Presentation order of the four blocks (each microphone and noise combination) was counterbalanced across the eight participants, and stimulus presentation levels were at least 10 dB above masked threshold.

Results and Discussion

Hit and false alarm rates were used to compute a sensitivity index (*d'*) for each subject in each condition. *d'* is monotonically related to proportion correct, but is not influenced by response bias. Table 2 lists percent correct values and approximate equivalent *d'* values. Because *d'* could not be calculated if were are either no hits (performance at floor) or no misses (performance at ceiling), if a participant had all hits and no false alarms in a particular condition (floor performance never occurred), the number of hits was reduced by one-half and the number of false alarms was increased by one-half before converting to proportion correct scores. This procedure resulted in a maximum *d'* score of 3.74 and allowed *d'* to be calculated for all conditions.

Table 2: Proportion correct values and equivalent *d'* values

Proportion correct	*d'* value
.97	3.74
.90	2.56
.80	1.68
.70	1.05
.60	.51
.50	0.00

Results were analyzed using a 2 × 2 × 6 repeated-measures ANOVA with microphone, noise, and consonant category serving as independent variables. Main effects were observed for all three factors (microphone type, noise type, and consonant type). DRT performance was significantly better with the boom microphone stimuli ($\underline{M}$ = 2.41) than with the throat microphone stimuli ($\underline{M}$ = 1.43), $\underline{F}(1, 9) = 225.99$, $p < .05$. As expected, performance was best in the no-noise condition ($\underline{M}$ = 2.73) compared to the 106 dB noise condition ($\underline{M}$ = 1.11), $\underline{F}(2, 18) = 127.65$, $\underline{p} < .05$. Inspection of the significant main effect of consonant type, $\underline{F}(5, 45) = 15.17$, $\underline{p} < .05$, showed that the voicing, sibilation, and compactness categories were perceived better than the other three consonant types.

Beyond the simple main effects, the significant three-way interaction ($\underline{F}(5,45) = 9.40$, $\underline{p} < .05$) demonstrated that speech intelligibility was influenced by microphone type, noise, and consonant category. These results are summarized in Figure 1, and the consonant category abbreviations can be found in Table 1.

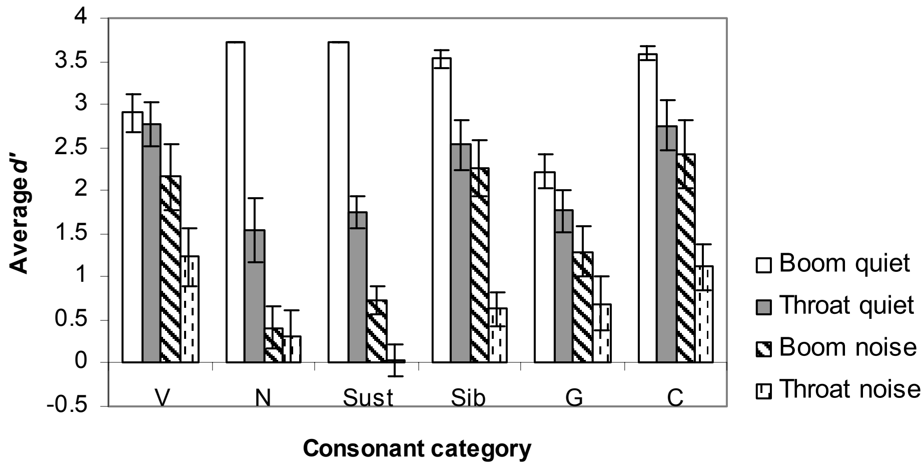

Figure 1: DRT performance as a function of microphone type, noise, and consonant category

The main effects are evident in Figure 1; the boom microphone almost always performs better than the throat microphone, and performance always declines significantly in noise (regardless of microphone type). Exceptions based on consonant category do occur. For example, performance between the boom and throat microphones was similar for the voicing category in quiet. This result is expected, as the throat microphone is in an ideal location to encode vibration, or lack thereof, of the vocal cords. Thus, information after the speech signal passes through the articulators is not essential for perception of voicing. Performance also was similar in the graveness category (in quiet and noise), which is dependent on frequency (see Table 1), but the distinctions involve lower frequencies that are transmitted by both microphones. In contrast, the absence of the effects of the articulators is rather evident for consonants with broadband frequency characteristics, such as /z/ of the sibilation category. The Appendix contains spectrographs of the words THEE and ZEE recorded in 106 dB(A) noise for the boom and throat microphones. Note the lack of distinction between the spectrograms of the throat microphone. This same principle holds for the compactness category, where distinctiveness of the consonants is dependent on perception of a broad or narrow frequency spectrum produced by the articulators, and the throat microphone fails to encode this information, particularly in noise. After inspection of the spectrograms, it is unclear why performance in noise dropped so low for both microphones in noise for the nasality and sustention categories.

The current results clearly demonstrate that while the throat microphone enhances the signal-to-noise ratio, the insensitivity to the effects of the articulators on the speech signal degrades speech intelligibility compared to the acoustic boom microphone. The DRT results indicated that some consonants are affected more than others by noise and/or use of a throat microphone. Spectrograms (see examples in the Appendix) revealed which acoustic features of the consonants were affected by microphone type and noise.

Because significant speech intelligibility differences occurred between acoustic and throat microphones, two other contact microphone systems were examined. It is possible that these bone-conduction systems that pick up vibrations from the entire vocal tract will perform better than the throat microphone which encodes information before the level of the articulators. In addition, signals transmitted through bone (a hard surface) may contain more information than those transmitted from the throat (soft tissue).

EXPERIMENT 2: BONE-CONDUCTION SYSTEMS

This Experiment used the same procedures as Experiment 1, except that the stimuli were recorded from head and ear bone-conduction communication systems. The head equipment consisted of a bone-conduction microphone on top of the skull (frequency response of approximately 200 Hz to 5000 Hz) and four bone-conduction speakers (frequency response of approximately 300 Hz to 3000 Hz) that contact the upper jaw bone and sides of the forehead on each side of the head (Temco Communications, Inc. model HG17). The ear equipment contained a bone-conduction microphone and an acoustic speaker contained in an earpiece that is placed in the ear canal and worn under a helmet (Temco Communications, Inc. model EM-P2). This microphone has a frequency response of approximately 200 Hz to 5000 Hz and the receiver 100 Hz to 3500 Hz. This particular model is designed to be used in noisy environments above 95 dB. Unfortunately, the person who recorded the boom and throat microphone stimuli was no longer available, so a different male speaker recorded the head and ear microphone stimuli. Whereas this situation is not ideal, it is highly unlikely that differences in the boom/throat and head/ear microphones are due to the recordings and not the microphones themselves. To support this notion, DRT data from an earlier study with a different speaker using the boom microphone were compared to the data of Experiment 1. Speech intelligibility was similar as for the speaker used in Experiment 1 (d' = 2.48 in the earlier study vs. d' = 2.41 in Experiment 1).

Thresholds for the two different speaker systems were measured using stimuli from the DRT task, and testing presentation levels were 10 dB above masked threshold. The head equipment prevented the helmet earcups from forming a tight seal, and therefore, participants had to wear earplugs in the noise condition in order to be protected from the 106 dB noise. The ear microphone/speaker system was worn in the right ear and one earplug was placed in the left ear during the noise condition. Eight participants, different from those in Experiment 1, completed the DRT task.

Results and Discussion

Results were compiled in the same manner as in Experiment 1, where hits and false alarms were used to calculate d'. Results were analyzed using a 2 × 2 × 6 repeated-measures ANOVA with microphone, noise, and consonant category serving as independent variables. Significant main effects occurred for all three independent variables. The head microphone and speaker combination ($\underline{M}$ = 1.648) performed better than the earpiece system ($\underline{M}$ = .976), $\underline{F}(1, 7) = 13.13$, $p < .05$, and performance was better in quiet ($\underline{M}$ = 1.74) than in noise ($\underline{M}$ = .879), $\underline{F}(1,7) = 27.59$, $p < .05$. Evaluation of the main effect of consonant category ($\underline{F}(5,35) = 3.54$, $p < .05$) revealed that intelligibility for the sibilation and sustention categories was less than for the other four categories.

Speech intelligibility is explained more completely by the microphone × consonant category interaction ($\underline{F}(5,35) = 9.24$, $p < .05$). As can be seen in Figure 2, performance was similar for the two systems in the nasality and sustention categories, but otherwise, intelligibility was better with the head microphone/speaker system. The spectrograms do not readily reveal why performance is similar in these two categories.

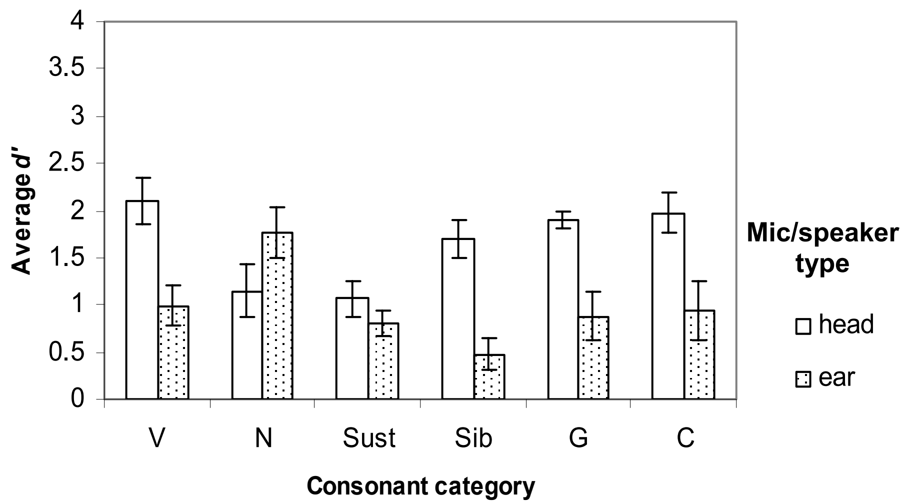

Figure 2: DRT performance as a function of microphone type and consonant category.

One advantage of *d'* as a sensitivity measure is that results between experiments legitimately can be compared. Figure 3 represents data combined over consonant category from Experiments 1 and 2, and it is evident that the boom microphone and HGU-56/P helmet earphones produce the best speech intelligibility, even in noise. The difference between the quiet and noise condition for the boom microphone may be artificially large because few errors were made, and even a few errors affect *d'* when performance is so good. The more important microphone comparison is in noise; the only system that had somewhat acceptable performance was the boom microphone at approximately 78 percent correct (performance probably would be even better with the Communication Earplug). Performance is at an acceptable level in quiet for the throat and head microphones, but is not acceptable in noise. Finally, the ear microphone/speaker results in the lowest performance even in quiet conditions (approximately 76 percent correct compared to 95 percent correct for the boom microphone in quiet). Whereas these effects are representative of microphone performance in general, keep in mind that both Experiments also exhibited interactions, where performance was affected by consonant category (see Figures 1 and 2).

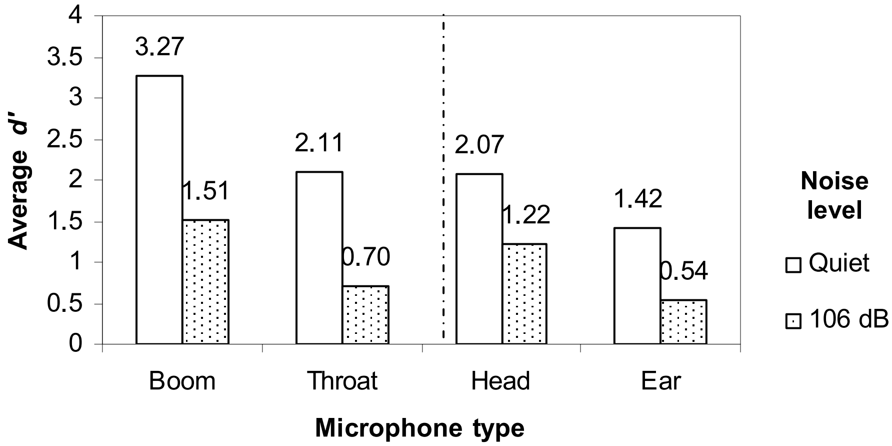

Figure 3: DRT performance as a function of microphone type and noise. Experiment 1 data are in the left panel and Experiment 2 data are in the right panel.

GENERAL DISCUSSION

The problem of noise and its detrimental effects on communication and hearing loss typically focuses on the "listener" (receiver). Whereas devices such as helmet earmuffs and the Communication Earplug can be useful (especially for hearing protection), speech intelligibility is still dependent on the quality of the original signal produced by the "talker" (sender). As noted above, the effectiveness of the noise-canceling boom microphone is reduced under various flying conditions that create unpredictable and highly variable noise. If this unpredictable ambient noise could be eliminated in the transmitted speech signal, the signal-to-noise ratio would be enhanced, and speech intelligibility also might be improved. Contact microphones greatly reduce or eliminate ambient noise because the microphone has a higher impedance that is matched only by vibrations on a surface (e.g., the skull) and not by the vibration of molecules in the air. Thus, the signal-to-noise ratio in a noisy environment is better than that of an acoustic microphone.

Even though the contact microphones have a better speech-to-noise ratio than the boom microphone, speech intelligibility was adversely affected by use of these microphones, particularly in noise. The most probable cause is that none of the contact microphones effectively encode information from the articulators, and this information is essential for differentiating consonants (e.g., the broadband noise in /z/ produced by the tongue). The results are troubling in that the DRT task represents a closed set of words and should be conditions where intelligibility is best (11). Whereas normal flight procedures also have standard communication phrases (basically a closed set), nonstandard speech will most likely occur in emergency or high-intensity combat situations. These are precisely the situations in which good speech intelligibility is critical, and the current results show that intelligibility using contact microphones is poorer than with the use of boom microphones even under the most benign of situations. Thus, it is recommended that contact microphones and speakers not be used in noisy environments where fast and accurate speech perception is critical.

ACKNOWLEDGEMENTS AND DISCLAIMER

We thank Mr. Christopher Madsen for his assistance supporting the software used to conduct this research, CPT Martin Robinette and SGT Kent Loo for their assistance in conducting this study, Ms. Janet Mauldin for her usual expert secretarial support, and our editor, Ms. Linda Burt. This research was funded in part by the U.S. Army Medical Research and Materiel Command In-house Laboratory Independent Review initiative.

The opinions, interpretations, conclusion, and recommendations are those of the authors and are not necessarily endorsed by the U.S. Army and/or the Department of Defense.

REFERENCES

[1] American National Standards Institute. Method for measuring the intelligibility of speech over communication systems. 1989; Melville, NY: Standards Secretariat, Acoustical Society of America. ANSI S3.2-1989 (R1999).

[2] Avaaz Innovations, Inc. ECoS Version 2™ Experiment generator and controller. North London, Ontario. 2002.

[3] Egan, J.P. Articulation testing methods. Laryngoscope 1948; 58(9): 955 – 991.

[4] Horie, S. Improvement of occupational noise-induced temporary threshold shift by active noise control earmuff and bone conduction microphone. J Occup Health 2002; 44(6): 414 – 420.

[5] Letowski, T. Bone conduction communication: Applications and limitations 2003. Paper presented at the National Hearing Conservation Association 28th Annual Conference, February 20-22, at Dallas, TX.

[6] Martin, D.W. Magnetic throat microphones of high sensitivity. J Acoust Soc Am 1947; 19(1): 40 – 43.

[7] Moser, H.M. and Dreher, J.J. Operational tests of miniature microphones and receivers. Columbus, OH: The Ohio State University Research Foundation 1956; Technical Report No. 36 for Contract No. AF 19(604)-1577.

[8] Nicolosi, L., Harryman, E., and Kresheck, J. Terminology of Communication Disorders in Speech-Language-Hearing (3rd Edition) 1989; Baltimore: Williams and Wilkins.

[9] Oyer, H.J. Relative intelligibility of speech recorded simultaneously at the ear and mouth. J Acoust Soc Am 1955; 27(6): 1207 – 1212.

[10] Snidecor, J.C., Rehman, I., and Washburn, D.D. Speech pickup by contact microphone at head and neck positions. J Speech Hear Res 1959; 2(3): 277 – 281.

[11] Suter, A.H. The effects of noise on speech and warning signals. Aberdeen Proving Ground, MD: U.S. Army Human Engineering Laboratory 1989; AD-A212 520.

APPENDIX

Spectrograms for stimuli recorded in 106 dB(A) noise with a throat and an acoustic microphone.

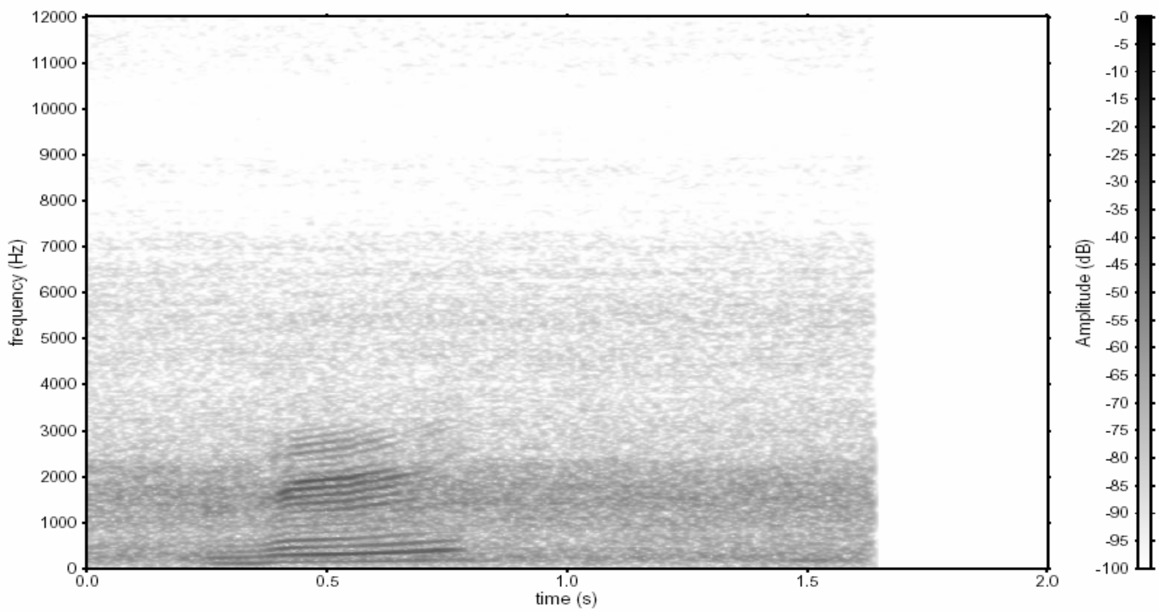

Figure A1. Spectrogram for THEE recorded with a throat microphone.

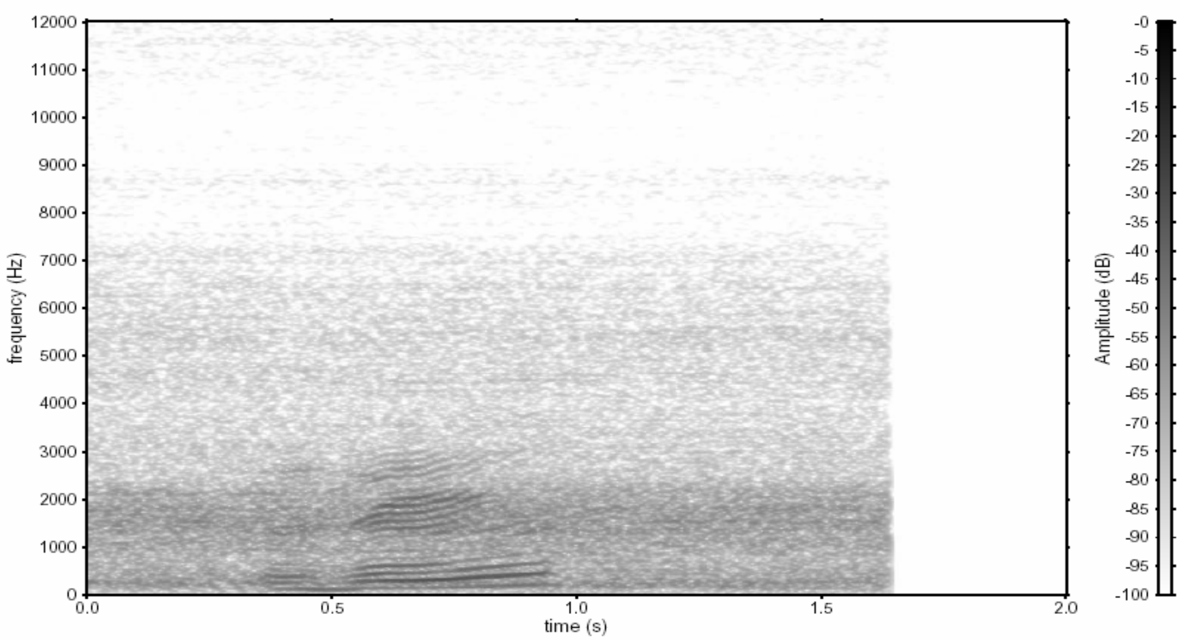

Figure A2. Spectrogram for ZEE recorded with a throat microphone.

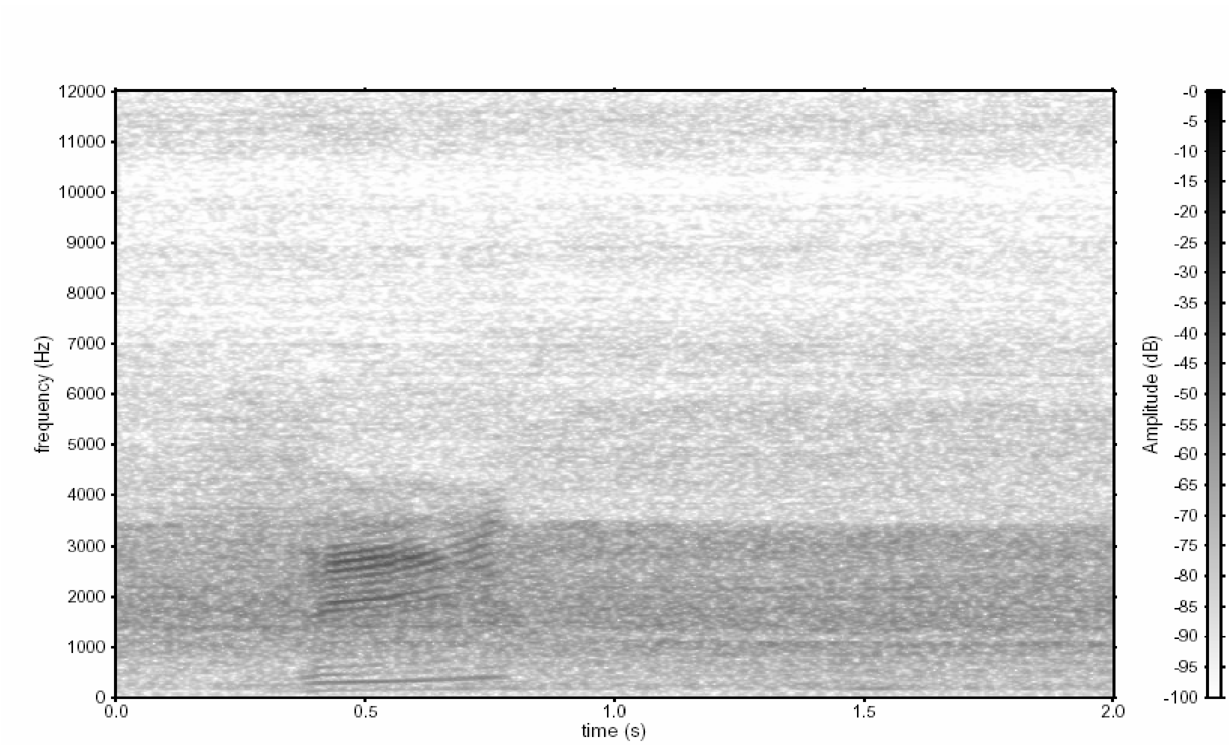

Figure A3. Spectrogram for THEE recorded with a boom microphone.

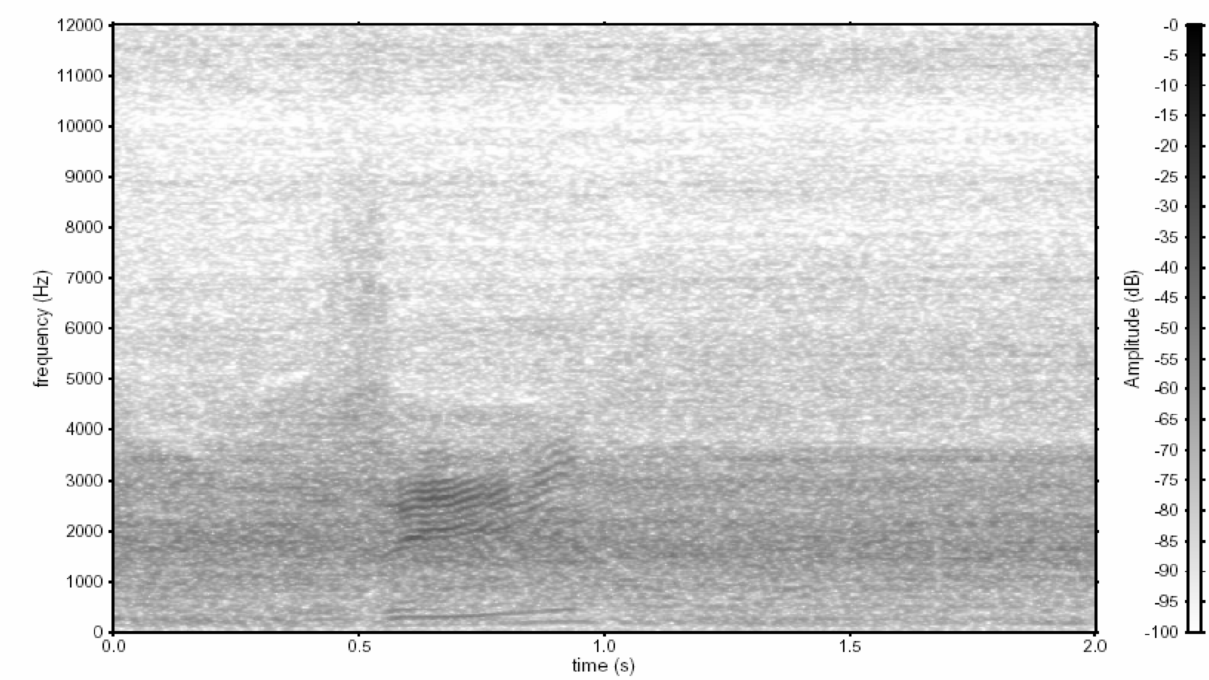

Figure A4. Spectrogram for ZEE recorded with a boom microphone.